

The Origin of Fermi Arcs in Cuprate Pseudogap States and Strong Constraints on Viable Theories of High-Temperature Superconductivity

Mike Guidry⁽¹⁾, Yang Sun⁽²⁾, and Cheng-Li Wu⁽³⁾

⁽¹⁾*Department of Physics and Astronomy, University of Tennessee, Knoxville, Tennessee 37996–1200, USA*

⁽²⁾*Department of Physics, Shanghai Jiao Tong University, Shanghai 200240, People's Republic of China*

⁽³⁾*Physics Department, Chung Yuan Christian University, Chung-Li, Taiwan 320, ROC*

(Dated: February 5, 2008)

A full Fermi surface exists in underdoped high-temperature superconductors if the temperature T lies above the pseudogap temperature T^* . Below T^* only arcs of Fermi surface survive, scaling with T/T^* as $T \rightarrow 0$, with T^* displaying strong doping dependence. There is no accepted explanation for this behavior. We show that generalizing the BCS theory of normal superconductivity to include d -wave pairs and antiferromagnetism leads to the origin and doping dependence of the T^* scale, and a quantitative description of Fermi arcs. These results place strong constraints on viable theories of high-temperature superconductivity.

High-temperature superconductors were discovered more than 20 years ago [1], but there is no uniformly-accepted explanation for their properties [2]. They exhibit pseudogaps (PG) at lower hole doping: partial energy gaps occurring below a temperature T^* but above the superconducting critical temperature T_c [3], with T^* and T_c exhibiting strong and opposite doping dependence for low hole-doping. It is widely believed that understanding the PG states and the T^* scale is central to understanding the high- T_c mechanism [2, 4].

For normal superconductors—described well by Bardeen–Cooper–Schrieffer (BCS) theory [5]—the Fermi surface (highest occupied fermion energy level) is key to understanding superconductivity. In optimally-doped and overdoped high- T_c compounds a “normal” Fermi surface exists, and BCS theory utilizing d -wave singlet hole pairs seems applicable [2]. In underdoped cuprates this picture fails: PG states (lying between T^* and T_c) have anomalous Fermi surfaces, with spectral strength pushed away from the expected Fermi energy by partial energy gaps. This is termed “gapping out”, or (more loosely) “loss”, of the Fermi surface.

Angle-resolved photoemission spectroscopy (ARPES) [6, 7] probes electronic properties [6, 8] of states in high-temperature superconductors. ARPES data suggest a decreased state density near the Fermi energy for temperature $T < T^*$ that is anisotropic in momentum, with strong T -dependence. A full Fermi surface is observed for $T > T^*$; as the temperature decreases below T^* , only arcs centered on the d -wave nodal lines (diagonals in momentum space) survive [9, 10, 11], scaling as T/T^* to zero length for $T \rightarrow 0$ [11]. Thus gapping is anisotropic in momentum space, with an ungapped Fermi surface surviving only in the form of temperature-dependent *Fermi arcs*.

Earlier theoretical work on Fermi arcs has employed a variety of approaches, including precursor pairs [12, 13], the perturbative renormalization group [14], high-temperature expansions in the t - J model [15], RVB models with strong gauge coupling [16], and time-reversal violating phases [17], but there is no agreed-upon explanation of Fermi arcs. The issue received renewed focus because of results by Kanigel, et al [11] that place the strongest constraints yet on the nature of Fermi arcs. We report here a quantitative description of the

Fermi-arc data in Ref. [11], with a theory that also consistently predicts the pseudogap temperature T^* at the heart of the Fermi-arc scaling behavior. Thus, we offer a comprehensive solution to one of the biggest mysteries in high-temperature superconductivity: the origin of the pseudogap and of Fermi arcs in underdoped cuprates. More importantly, our results imply that only high- T_c models with a strongly-constrained set of properties can be consistent with the detailed attributes of Fermi arcs and the scale T^* .

In earlier work we proposed an SU(4) model for the ground state properties of cuprate systems. Generalized SU(4) coherent states are the simplest Hartree–Fock–Bogoliubov variational solutions that incorporate antiferromagnetism (AF) competing with d -wave superconductivity on a lattice with no double occupancy [18, 19, 20, 21]. They generalize BCS to incorporate AF self-consistently. If AF correlations vanish, the SU(4) gap equations reduce to the BCS gap equations; if instead singlet pairing vanishes, the SU(4) gap equations describe an AF spin system with Néel order; for finite AF and SC correlations, the SU(4) gap equations describe the evolution with doping (P) and temperature (T) of d -wave Cooper pairs interacting with strong AF correlations.

SU(4) states forbid double occupancy by symmetry, not by projection [19]. The exact zero- T ground state at half filling is an AF Mott insulator that evolves rapidly with hole doping into a superconductor with strong AF correlations in the underdoped region, and finally into a singlet, d -wave superconductor beyond a critical hole doping $P_q \simeq 0.16\text{--}0.19$ [20]. At finite T for $P > P_q$, pairing is reduced by thermal fluctuations and the pairing gap vanishes at T_c . In the underdoped region ($P < P_q$) there are quantum fluctuations associated with AF correlations in addition to thermal fluctuations at finite T . The interplay between AF, pairing, and thermal fluctuations produces PG states above T_c in which the fermionic pairs interact through AF correlations. Thus the SU(4) model identifies the energy gap Δ_q opened by the AF correlations as the pseudogap, while the (singlet and triplet fermion) pairs in these states may be viewed as d -wave preformed pairs.

The SU(4) model can reproduce both the pseudogap and pairing gap with doping dependence quantitatively in agreement with data (see Ref. [21]). The SC transition tempera-

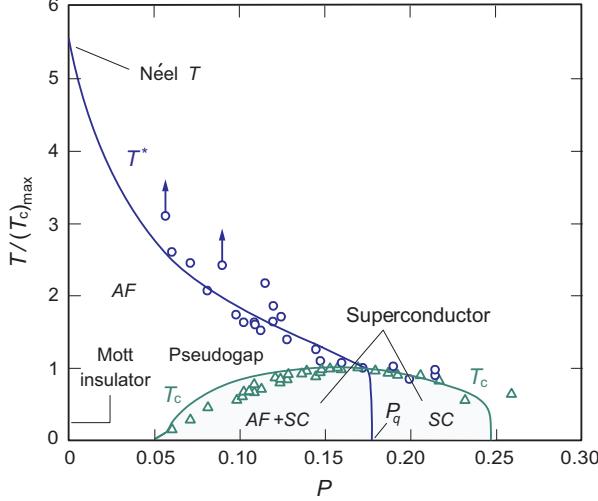


FIG. 1: (Color online) SU(4) cuprate phase diagram compared with data. Strengths of the AF and singlet pairing correlations were determined by global fits to cuprate data [20, 21]. The PG temperature is T^* and the SC transition temperature is T_c . The AF correlations vanish, leaving a pure singlet d -wave condensate, above the critical doping P_q . Dominant correlations in each region are indicated by italic labels. Data from Ref. [23] (arrows indicate lower limits).

tures T_c and PG transition temperatures T^* are also well reproduced, as illustrated in Fig. 1. This is our first significant finding: SU(4) coherent states give descriptions of pairing gaps and pseudogaps, and their transition temperatures T_c and T^* , that are in quantitative agreement with cuprate data.

The 15 generators $\{\mathcal{G}_i\}$ of SU(4) are $M = \frac{1}{2}(n - \Omega)$ and

$$\begin{aligned} D^\dagger &= \sum_{\mathbf{k}} s_{gk} c_{\mathbf{k}\uparrow}^\dagger c_{-\mathbf{k}\downarrow}^\dagger & \pi_{ij}^\dagger &= \sum_{\mathbf{k}} s_{gk} c_{\mathbf{k}+\mathbf{q},i}^\dagger c_{-\mathbf{k},j}^\dagger \\ Q_{ij} &= \sum_{\mathbf{k}} c_{\mathbf{k}+\mathbf{q},i}^\dagger c_{\mathbf{k},j} & S_{ij} &= \sum_{\mathbf{k}} c_{\mathbf{k},i}^\dagger c_{\mathbf{k},j} - \frac{1}{2}\Omega\delta_{ij} \end{aligned} \quad (1)$$

and hermitian conjugates, where \uparrow, \downarrow, i , and j are spin indices, $\mathbf{q} \equiv (\pi, \pi, \pi)$, the maximum number of doped holes or particles that can form coherent pairs (assuming the half-filled normal state as vacuum) is Ω , and s_{gk} is the algebraic sign of the d -wave pair formfactor

$$g(\mathbf{k}) = g(k_x, k_y) = \cos k_x - \cos k_y. \quad (2)$$

In these expressions M is the charge with $n = \sum_{\mathbf{k}} n_{\mathbf{k}}$ the electron number operator, and π_{ij}^\dagger , Q_{ij} , and S_{ij} are tensor components of the triplet pair, staggered magnetization, and spin operators $\vec{\pi}$, \vec{Q} and \vec{S} , respectively.

These generators are summed over momentum \mathbf{k} . They are appropriate for describing data that are not momentum-selected (for example, Fig. 1). However, ARPES Fermi-arc data exhibit explicit dependence on \mathbf{k} . The SU(4) formalism may be extended to deal with this case by viewing the individual \mathbf{k} components of the operators defined in Eq. (1) as the symmetry generators $\mathcal{G}(\mathbf{k})$:

$$\{\mathcal{G}(\mathbf{k})\} \equiv \{D^\dagger(\mathbf{k}), D(\mathbf{k}), \vec{\pi}^\dagger(\mathbf{k}), \vec{\pi}(\mathbf{k}), \vec{Q}(\mathbf{k}), \vec{S}(\mathbf{k}), n_{\mathbf{k}}\}$$

where, for example, the singlet pair generator D^\dagger in Eq. (1) is related to the \mathbf{k} -dependent generators $D^\dagger(\mathbf{k})$ by $D^\dagger = \sum_{\mathbf{k}>0} D^\dagger(\mathbf{k})$, and $\mathbf{k} > 0$ means either $k_x > 0$ or $k_y > 0$.

Instead of the global SU(4) symmetry generated by Eq. (1), the symmetry now is a direct product of \mathbf{k} -dependent SU(4) groups, $\prod_{\mathbf{k}>0} \otimes \text{SU}(4)_{\mathbf{k}}$. The corresponding Hamiltonian is

$$H = \sum_{\mathbf{k}>0} \varepsilon_{\mathbf{k}} n_{\mathbf{k}} - \sum_{\mathbf{k}, \mathbf{k}'>0} \{\chi_{\mathbf{k}\mathbf{k}'} \vec{Q}(\mathbf{k}) \cdot \vec{Q}(\mathbf{k}') + G_{\mathbf{k}\mathbf{k}'}^{(0)} D^\dagger(\mathbf{k}) D(\mathbf{k}') + G_{\mathbf{k}\mathbf{k}'}^{(1)} \vec{\pi}^\dagger(\mathbf{k}) \cdot \vec{\pi}(\mathbf{k}')\}, \quad (3)$$

where $\varepsilon_{\mathbf{k}}$ and $n_{\mathbf{k}}$ are single-particle energies and occupation numbers, respectively, and the interaction strengths are

$$\chi_{\mathbf{k}\mathbf{k}'} = \chi^0 |g(\mathbf{k})g(\mathbf{k}')| \quad G_{\mathbf{k}\mathbf{k}'}^{(i)} = G^{(i)} \delta_{\mathbf{k}} \delta_{\mathbf{k}'} \quad (i = 0, 1),$$

where $\delta_{\mathbf{k}} \equiv \delta(\theta)|g_0(\tilde{\mathbf{k}})|$. The pair formfactor is expressed in terms of the magnitude of the hole momentum $\tilde{\mathbf{k}}$ and its azimuthal angle θ : $g(\mathbf{k}) \equiv g(k_x, k_y) = g(\tilde{\mathbf{k}}, \theta)$, where $|g_0(\tilde{\mathbf{k}})|$ is $|g(\tilde{\mathbf{k}}, \theta)|$ at the antinodes ($\theta = 0, \pi/2$), where it is maximal, and $\delta(\theta) = 1$ except in a very narrow region near the nodal point, where it drops rapidly to zero at the node.

We shall term this \mathbf{k} -dependent formalism the $\text{SU}(4)_{\mathbf{k}}$ model. All results of the original \mathbf{k} -independent SU(4) model of Refs. [18, 19, 20, 21] (including those of Fig. 1) are recovered as a special case of the $\text{SU}(4)_{\mathbf{k}}$ model for observables that are dominated by contributions from near the Fermi surface ($\tilde{\mathbf{k}} = \mathbf{k}_f$) and averaged over \mathbf{k} directions. However, general solutions of the $\text{SU}(4)_{\mathbf{k}}$ model give new \mathbf{k} -anisotropic properties. Of direct relevance to the present discussion is that the temperature for the PG closure $T^*(\mathbf{k})$ becomes anisotropic in \mathbf{k} ; specifically, we derive in the $\text{SU}(4)_{\mathbf{k}}$ model

$$T^*(\mathbf{k}) \equiv T^*(k_f, \theta) = T^* |g(k_f, \theta)/g_0(k_f)| \quad (4)$$

where T^* is the gap closure temperature measured by ARPES along the antinodal ($\theta = 0, \pi/2$) direction, which is the maximum possible value of T^* and should be somewhat larger than T^* inferred from experiments that average over \mathbf{k} . See Ref. [22] for details. Equation (4) defines the full temperature and doping dependence of Fermi arcs. The ultimate physical reason for this result is that the singlet and triplet pairs interacting by AF interactions in the SU(4) PG state each carry a $g(\mathbf{k})$ formfactor, which introduces a \mathbf{k} dependence in the effective AF coupling and thus in T^* .

In Fig. 2(a) we show the behavior of $g(\mathbf{k})$ as a function of (k_x, k_y) . The components (k_x, k_y) or (k_f, θ) are constrained by the Fermi surface ($\tilde{\mathbf{k}}^2 = k_f^2$). Assuming an isotropic hole surface [dashed gray lines in Fig. 2(a)], we have

$$(\pi - k_x)^2 + (\pi - k_y)^2 = k_f^2 = 2\pi(1 + P) \quad (5)$$

(we consider a more realistic Fermi surface below). For a given doping P (thus k_f) and temperature T , with $T^*(\mathbf{k}) = T$ by virtue of Eq. (4) under the constraint (5), we obtain the angles θ_1 and θ_2 at which the PG closes (the angle θ is defined in the inset to Fig. 2(b)), and the length of the surviving Fermi arc is $k_f|\theta_2 - \theta_1| \equiv k_f \Delta\theta$.

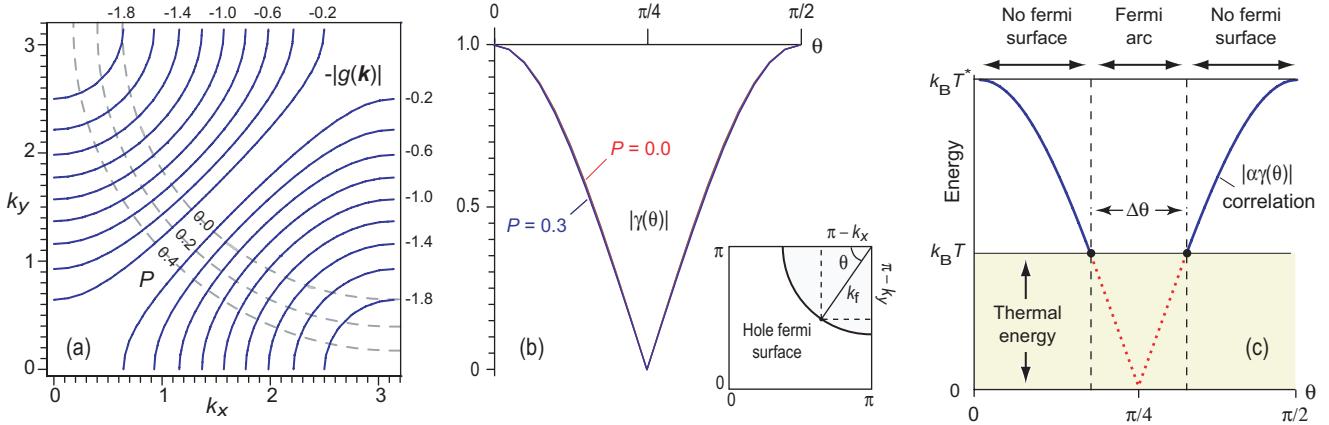


FIG. 2: (Color online) The \mathbf{k} -anisotropic factor, pseudogap correlation energy, and graphical Fermi-arc solution. (a) The \mathbf{k} -anisotropic factor $|g(\mathbf{k})|$ (blue lines) and contours of equal hole doping P (dashed gray lines), in the first Brillouin zone. (b) Correlation energy $\alpha\gamma(\theta)$ for unit $\alpha = kT^*$ evaluated along the Fermi surface [curves of constant P in part (a)] as a function of the angle θ (defined in inset). Curves corresponding to doping $P = 0\text{--}0.3$ lie almost on top of each other, indicating that $\gamma(\theta)$ is largely insensitive to doping. (c) Graphical solution of the Fermi-arc problem. The curve defines the PG correlation energy and horizontal lines correspond to constant thermal energy scales $k_B T$ associated with a given temperature T . Their intersections (black dots) represent points in the momentum space where the pseudogap is just closed by thermal fluctuations; these bracket arcs $\Delta\theta$ of surviving Fermi surface.

This result may be interpreted graphically. In Fig. 2(b) we show $\gamma(\theta) = |g(k_f, \theta)/g_0(k_f)|$ versus angle θ along different Fermi surfaces [dashed gray lines in Fig. 2(a)]. We see that $\gamma(\theta)$ is almost independent of doping P . Thus, solving Eq. (4) with the Fermi surface constraint (5) is equivalent to solving

$$k_B T^*(\mathbf{k}) = \alpha\gamma(\theta) \quad \alpha = k_B T^*, \quad (6)$$

with k_B the Boltzmann constant. The PG correlation energy $\alpha\gamma(\theta)$ depends on the \mathbf{k} direction θ [Fig. 2(c)]. The pseudogap closes when the thermal excitation energy $k_B T$ is comparable to the PG correlation energy. The intersections of horizontal lines of fixed $k_B T$ with the correlation energy curve [black dots in Fig. 2(c)] define Fermi-arc solutions θ_1 and θ_2 . Outside those points (solid blue curve), the correlation energy is larger than the energy of thermal fluctuations (shaded region), the PG opens, and the Fermi surface is gapped. Inside these points (dotted red curve), the thermal energy exceeds the correlation energy, the PG closes, and the Fermi surface exists in an arc $\Delta\theta$ between the dots. When $T > T^*$, the PG is closed in all directions and there is a full Fermi surface since $k_B T > \alpha$, and α is the maximum PG correlation energy.

The arc solution exemplified graphically in Figs. 2(b)–(c), or algebraically in Eq. (4), is our second significant finding: For given temperature T , the anisotropy of $\gamma(\theta)$ partitions the k -space uniquely into regions having a Fermi surface [$T > T^*(k_f, \theta)$] and those that do not [$T < T^*(k_f, \theta)$]. As Fig. 2(c) suggests, arc lengths decrease with decreasing T at fixed doping, with a full Fermi surface at $T = T^*$, but only the nodal points at $T = 0$. Doping dependence enters primarily through the maximum PG temperature T^* and the Fermi momentum k_f ; the temperature dependence enters through T/T^* . For fixed T , arcs shrink toward the nodal points with decreased doping because T^* in T/T^* increases at smaller doping (Fig.

1). However, if Fermi-arc length is measured by the fraction $\Delta\theta/(\pi/2)$ and the temperature is scaled by T^* , the weak doping dependence of $\gamma(\theta)$ ensures that $\Delta\theta/(\pi/2)$ versus T/T^* is almost independent of doping and compound.

Kanigel, et al [11] report fractional arc lengths versus T/T^* for Bi2212 that we plot in Fig. 3. Our theoretical solution for the fractional arc lengths [obtained from Fig. 2(c) or Eq. (4)] is the solid line in Fig. 3. Agreement between data and theory is remarkable, given that the theoretical curve has no adjustable parameters (it is determined completely by the parameters fixed previously in Fig. 1) and that we *predict* the scale T^* implicit in the data with the same theory (Fig. 1). Note that the different behavior at high and low temperatures in Fig. 3 (rapid drop in arc length for decreasing $T \sim T^*$, shifting to approximately linear decrease to zero arc-length at $T = 0$), is explained entirely by geometry in Fig. 2(c).

Realistic cuprate Fermi surfaces are flat near zone boundaries. The dashed curve in Fig. 3 uses the flatter Fermi surface shown inset lower right. The similarity of dashed and solid curves indicates that arc solutions depend only weakly on differing curvature in nodal and antinodal regions. Absolute arc lengths depend on doping and temperature. However, the scaled arc lengths of Fig. 3 are *near-universal functions only of the ratio T/T^** , largely independent of compound, doping, and Fermi surface details, as data suggest.

Figure 3 illustrates our third significant finding: agreement between theory and data with no parameter adjustment suggests that our SU(4) formalism outlines a complete solution of the Fermi-arc mystery. We now argue that this has important implications for understanding pseudogap behavior.

The data in Fig. 3 have been interpreted [11] as representing a rapid drop from a full Fermi surface at $T = T^*$, destroying the antinodal Fermi surface at essentially constant tempera-

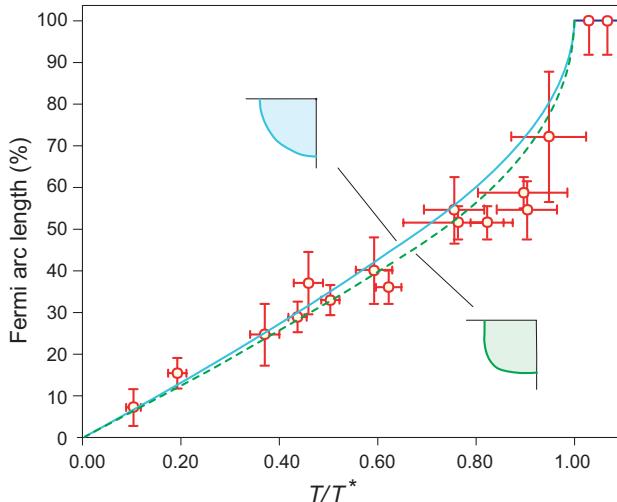


FIG. 3: (Color online) Fermi arc length versus temperature. Experimental arc length is displayed as a percentage of full Fermi surface length vs. T/T^* for underdoped Bi2212 [11]; the solid curve is our prediction for an isotropic Fermi surface (inset upper left); the dashed curve assumes a Fermi surface with flat antinodal segments (inset lower right). No parameters were adjusted to these data in either calculation and the curves are almost independent of doping.

ture, followed by a linear decrease of Fermi-arc length with decreasing T/T^* on the near-nodal region of the Fermi surface, extrapolating to zero arc length at the nodal points for $T/T^* \rightarrow 0$. This suggests that nodal and antinodal Fermi surfaces may be removed differently as T is lowered (see Ref. [24]). For example, it has been argued [11] that abrupt removal of the antinodal surface at $T \simeq T^*$ may be associated with a lattice vector connecting antinodal surfaces (Ref. [24]) speculates that gapping of the antinodal surface may be associated with charge ordering), while the linearly-varying removal of the nodal regions extrapolating to a nodal-point surface at $T = 0$ may support a nodal liquid picture [25].

The present results invite simpler hypotheses. Our unified analysis suggests that Fig. 3 is consistent with removal of both nodal and antinodal surfaces by the *same mechanism*. Therefore, the qualitatively different variation of arc length with T/T^* in nodal and antinodal regions is not sufficient to indicate that loss of Fermi surface proceeds by different mechanisms in these regions. Nor do the data of Ref. [11] imply unique support for nodal liquids. We have demonstrated specifically that the SU(4) model accounts quantitatively for Fermi arcs and T^* , without invoking nodal liquids.

We have shown that a theory of Fermi arcs must make two correct predictions: (1) the scale T^* and its doping dependence, and (2) that $T^*(\mathbf{k}) \propto \gamma(\theta)$. Any theory having a $T^*(\mathbf{k})$ consistent with Eq. (6) can describe the scaled data of Fig. 3, if T^* is taken from data. Hence, our fourth significant finding: scaled ARPES data (Fig. 3) test whether a pseudogap has a nodal structure similar to that of the superconducting state. But discriminating among different theories meeting this con-

dition requires more: a quantitative, self-consistent description of Fermi-surface loss and of the PG temperature scale T^* and its doping dependence (Fig. 1). We conclude that only a highly-restricted set of models can be consistent with the aggregate properties of Fermi arcs. Finally, although our results indicate that an SU(4) mean field can account for Fermi arcs, some properties of the pseudogap state are expected to be influenced significantly by SU(4) quantum fluctuations that we shall address in forthcoming papers.

We thank Elbio Dagotto, Pengcheng Dai, and Takeshi Egami for extensive discussion, and Elbio Dagotto for calling our attention to an error in our original formulation.

-
- [1] J. G. Bednorz and K. A. Müller, Z. Phys. **B 64**, 189 (1986).
 - [2] D. A. Bonn, Nature Phys. **2**, 159 (2006), and refs. therein.
 - [3] T. Timusk, and B. Statt, Rep. Prog. Phys. **62**, 61 (1999).
 - [4] M. Norman, D. Pines, and C. Kollin, Adv. Phys. **54**, 715 (2005).
 - [5] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **108**, 1175 (1957).
 - [6] M. R. Norman and C. Pépin, Rep. Prog. Phys. **66**, 1547 (2003).
 - [7] A. Damascelli, A. Hussain, and Z.-X. Shen, Rev. Mod. Phys. **75**, 473 (2003).
 - [8] E. Dagotto, Rev. Mod. Phys. **66**, 763 (1994).
 - [9] M. R. Norman, et al, Nature **392**, 157 (1998).
 - [10] X.-J. Zhou, et al, Phys. Rev. Lett. **92**, 187001 (2004).
 - [11] A. Kanigel, et al, Nature Phys. **2**, 447 (2006).
 - [12] J. R. Engelbrecht, A. Nazarenko, M. Randeria, and E. Dagotto, Phys. Rev. **B 57**, 13406 (1998).
 - [13] G. Preosti, Y. M. Vilk, and M. R. Norman, Phys. Rev. **B 59**, 1474 (1999).
 - [14] N. Furukawa, T. M. Rice, and M. Salmhofer, Phys. Rev. Lett. **81**, 3195 (1998).
 - [15] W. O. Putikka, M. U. Luchini, and R. R. P. Singh, Phys. Rev. Lett. **81**, 2966 (1998).
 - [16] X. G. Wen, and P. A. Lee, Phys. Rev. Lett. **80**, 2193 (1998).
 - [17] C. M. Varma, and L. Zhu, arXiv:cond-mat/0607777.
 - [18] M. W. Guidry, Rev. Mex. Fís. **45 S2**, 132 (1999); M. W. Guidry, L.-A. Wu, Y. Sun, and C.-L. Wu, Phys. Rev. **B 63**, 134516 (2001); L.-A. Wu, M. W. Guidry, Y. Sun, and C. L. Wu, Phys. Rev. **B 67**, 014515 (2003).
 - [19] M. W. Guidry, Y. Sun, and C.-L. Wu, Phys. Rev. **B 70**, 184501 (2004).
 - [20] Y. Sun, M. W. Guidry, and C.-L. Wu, Phys. Rev. **B 73**, 134519 (2006).
 - [21] Y. Sun, M. W. Guidry, and C.-L. Wu, Phys. Rev. **B 75**, 134511 (2007).
 - [22] Y. Sun, M. W. Guidry, and C.-L. Wu, arXiv:0705.0818
 - [23] P. Dai, et al, Science **284**, 1344 (1999); M. Oda, et al, Physica **C 281**, 135 (1997); T. Watanabe, et al, Phys. Rev. Lett. **84**, 5848 (2000); J. C. Campuzano, et al, Phys. Rev. Lett. **83**, 3709 (1999); J. Wuyts, V. V. Moshchalkov, and Y. Bruynseraede, Phys. Rev. **B 53**, 9418 (1996); M. Takigawa, et al, Phys. Rev. **B 43**, 247 (1991); T. Ito, K. Takenaka, and S. Uchida, Phys. Rev. Lett. **70**, 3995 (1993).
 - [24] K. McElroy, Nature Phys. **2**, 441 (2006).
 - [25] L. Balents, M. P. A. Fisher, and C. Nayak, Int. J. Mod. Phys. **B 12**, 1033 (1998).